CRC 615 / March 1992

Polynomial Families of Distributions

D. R. Divgi

50 Years CNA 1992

CENTER FOR NAVAL ANALYSES

4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268

Work conducted under contract N00014-91-C-0002.

This Research Contribution represents the best opinion of CNA at the time of issue. It does not necessarily represent the opinion of the Department of the Navy.

REPORT DOCUMENTATION PAGE

Form Approved
OPM No. 0704-0188

1. AGENCY USE ONLY *(Leave Blank)*	2. REPORT DATE March 1992	3. REPORT TYPE AND DATES COVERED Final

4. TITLE AND SUBTITLE

Polynomial Families of Distributions

5. FUNDING NUMBERS

C - N00014-91-C-0002

PE - 65153M

PR - C0031

6. AUTHOR(S)

D.R. Divgi

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Center for Naval Analyses
4401 Ford Avenue
Alexandria, Virginia 22302-0268

8. PERFORMING ORGANIZATION REPORT NUMBER

CRC 615

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Commanding General
Marine Corps Combat Development Command (WF 13F)
Studies and Analyses Branch
Quantico, Virginia 22134

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

It is often necessary to estimates the population distribution of a random variate from a sample of observed values. Standard parametric families may not provide satisfactory fit to the data. A polynomial family is constructed by assuming that the distribution function G is a constrained polynomial of the cumulative distribution F of a convenient parametric family. Polynomial families offer great flexibility in data fitting, while retaining the important feature of parametric families that information in the data is condensed into a moderate number of values.

14. *SUBJECT TERMS*

CH1 square test, Maximum likelihood estimation, Numerical analysis, Polynomials, Statistical analysis, Statistical distributions

15. *NUMBER OF PAGES*

18

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT CPR	18. SECURITY CLASSIFICATION OF THIS PAGE CPR	19. SECURITY CLASSIFICATION OF ABSTRACT CPR	20. LIMITATION OF ABSTRACT SAR

50 Years CNA 1992 CENTER FOR NAVAL ANALYSES

4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268 • (703) 824-2000

7 April 1992

MEMORANDUM FOR DISTRIBUTION LIST

Subj: CNA Research Contribution 615

Encl: (2) CNA Research Contribution 615, *Polynomial Families of Distributions*, by D.R. Divgi, March 1992

1. Enclosure (1) is forwarded as a matter of possible interest.

2. It is often necessary to estimate the population distribution of a random variate from a sample of observed values. Standard parametric families may not provide satisfactory fit to the data. A polynomial family is constructed by assuming that the distribution function G is a constrained polynomial of the cumulative distribution F of a convenient parametric family. Polynomial families offer great flexibility in data fitting, while retaining the important feature of parametric families, namely, that information in the data is condensed into a moderate number of values.

3. Research contributions are distributed for their potential value in other studies and analyses. They do not necessarily represent the opinion of the Department of the Navy.

Jamil Nakhleh
Director
Operations and Support Division

Distribution List:
Reverse page

Subj: Center for Naval Analyses Research Contribution 615

Distribution

SNDL

A1	DASN - MANPOWER
A1H	ASSTSECNAV MRA
A2A	CNR
A5	PERS-11B
A5	PERS-23
A6	HQMC MPR & RA
	Attn: Code M
	Attn: Code MP
	Attn: Code MR
	Attn: Code MA
	Attn: Code MPP-54
FF38	USNA
	Attn: Nimitz Library
FF42	NAVPGSCOL
FF44	NAVWARCOL
	Attn: E-111
FJA1	COMNAVMILPERSCOM
FJA13	NAVPERSRANDCEN
	Attn: Technical Director (Code 01)
	Attn: Dir., Testing Systems (Code 13)
	Attn: Technical Library
	Attn: Dir., Personnel Systems (Code 12)
	Attn: CAT/ASVAB PMO
	Attn: Manpower Systems (Code 11)
FJB1	COMNAVCRUITCOM
FT1	CNET
V12	CG MAGTEC
V12	CG MCCDC
	Attn: Commanding General
	Attn: Warfighting Center

OTHER

Military Accession Policy Working Group (18 copies)
Defense Advisory Committee on Military Personnel Testing (8 copies)

CRC 615 / March 1992

Polynomial Families of Distributions

D. R. Divgi

Operations and Support Division

50 Years
CNA 1992
CENTER FOR NAVAL ANALYSES
4401 Ford Avenue • Post Office Box 16268 • Alexandria, Virginia 22302-0268

ABSTRACT

It is often necessary to estimate the population distribution of a random variate from a sample of observed values. Standard parametric families may not provide satisfactory fit to the data. A polynomial family is constructed by assuming that the distribution function G is a constrained polynomial of the cumulative distribution F of a convenient parametric family. Polynomial families offer great flexibility in data fitting, while retaining the important feature of parametric families that information in the data is condensed into a moderate number of values.

THIS PAGE INTENTIONALLY LEFT BLANK

EXECUTIVE SUMMARY

The Armed Services Vocational Aptitude Battery consists of multiple-choice tests. New kinds of computerized tests are being developed and evaluated. Distributions of scores on these new tests can be very different from those of scores on multiple-choice tests. The same is true of the computerized adaptive version of the ASVAB. Distributions observed in raw data contain sample error. Smoothing of these distributions is useful in reducing the errors in statistical analyses, and also in displaying the distributions.

The actual score distribution may not belong to any of the familiar families of distributions. In such a case, one can begin with a suitable family and then generalize it. In the generalized family, the cumulative distribution function G is a polynomial of F, the distribution function of the original family. This approach can generate distributions with a wide variety of shapes. This research contribution presents some theory of such general families of distributions.

THIS PAGE INTENTIONALLY LEFT BLANK

CONTENTS

INTRODUCTION

It is often necessary to estimate the distribution of a random variate X from a sample of observations. Standard parameteric families may not provide satisfactory fit to the data. For example, the distribution may be multimodal. We can use nonparametric density estimation, but then we lose the convenience of summarizing the information in the data in a moderate number of values. It would be useful to achieve a compromise between standard parametric families and nonparametric methods. This can be done by defining a family in which the number of parameters can be increased indefinitely until satisfactory fit to the data is obtained.

POLYNOMIAL FAMILIES

Let $F(x, \underline{\theta})$ be any parametric cumulative distribution function. (The underscore in $\underline{\theta}$ indicates that it is a vector.) Let $G(x, \underline{\theta}, \mathbf{a})$ be a polynomial of F with the form

$$G = F + \sum_{k=1}^{p} a_k g_k(F) \quad , \tag{1}$$

where function g_k is a polynomial of degree $k + 1$ and contains the factor $F(1 - F)$. The coefficients **a** are such that G is monotone nondecreasing in (0,1). The factor $F(1 - F)$ ensures that $G = 0$ when $F = 0$, and $G = 1$ when $F = 1$. Thus G, too, is a cumulative distribution function and hence can be used for fitting observed data. The functions can be of the simple form

$$g_k(F) = F(1 - F)\, F^{k-1} \quad , \tag{2}$$

but then the polynomial in equation 1 contains successive terms that are strongly correlated, which can lead to ill-conditioned matrices and numerical instabilities while estimating the coefficients. Better expressions for these functions are given in appendix A. The distribution F, which is a special case of G, will be referred to as the "base" distribution. The expressions are a compromise between simplicity and spreading out the zeroes of the polynomials. If ninth degree polynomials fail to yield a good fit, one should probably try a different base distribution or try transforming the data.

In principle, the base distribution can have any form with any number of parameters. In practice, its choice depends on ease of computing F and its derivatives and on its suitability for the data in hand. The normal distribution is a natural choice if x can take any real value. If x can take only positive values, the Weibull is more convenient than the gamma distribution. The beta distribution is a natural (although not convenient) choice if x has a known finite range. The variable X can be discrete as well as continuous. Then, depending on the nature of X, the distribution may be binomial, hypergeometric, Poisson, and so on. Appendix B provides formulas for fitting distributions of scores on multiple-choice tests, using the negative hypergeometric as the base distribution.

In data fitting, the value of p may be set a priori or determined in a stepwise fashion. In the latter case, we begin by fitting F. Then, for each succesive value of p, we reestimate all parameters (including those of F) and decide whether addition of the last term provides a significant improvement in fit to the data.

Polynomial families are useful because they are extremely flexible. Given the freedom in choosing the base distribution as well as the degree of the polynomial, a wide variety of shapes can be obtained. In the illustration presented later in this report, a Weibull base and only two coefficients in the polynomial provide excellent fit to a bimodal distribution.

MAXIMUM LIKELIHOOD ESTIMATION

Maximum likelihood estimation (MLE) has optimal asymptotic properties. In use with polynomial families, MLE can create a problem in computation. In iterative fitting, coefficients **a** may be such that the polynomial for G is not monotone at each observed value of x. As a result, the density may be negative or zero; hence, its logarithm may not exist at some values of x. The algorithm to calculate log likelihood must check for this possibility, and the maximization routine must contain steps to deal with the problem if it arises. Appendix A gives equations for MLE by the Newton-Raphson procedure.

MINIMUM CHI-SQUARE ESTIMATION

Computations are substantially simpler if, instead of maximum likelihood, we use minimum chi-square with the objective function defined as follows. Let N be the sample size, $m \leq N + 1$ a positive integer, $x_{(i)}$, $i=1, 2, \ldots, N$ the order statistics, and

$$0 = \lambda_0 < \lambda_1 < \lambda_2 \ldots . < \lambda_{m-1} < \lambda_m = 1 \quad . \tag{3}$$

Let $N_h = [(N + 1)\lambda_h] + 1$ where $[y]$ is the largest integer less than y, and $n_h = N_h - N_{h-1}$. The λ's must be such that each $n_h > 0$. By definition, $x_{(0)}$ and $x_{(m)}$ are smallest and largest possible values of x so that $G(x_{(0)}, \underline{\theta}, a) = 0$ and $G(x_{(m)}, \underline{\theta}, a) = 1$. The quantity to be minimized is

$$Q = \sum_{h=1}^{m} [(N + 1)\{G(x_{(N_h)}, \underline{\theta}, a) - G(x_{(N_{h-1})}, \underline{\theta}, a)\} - n_h]^2/n_h \quad . \tag{4}$$

Although there is no rule for choosing λ's, it seems desirable to space them uniformly so that Q is equally sensitive to different parts of the distribution.

If the parameters $\underline{\theta}$ and $\underline{a}$ are known, the quantity in curly brackets is a spacing of order n_h, with expected value $n_h/(N + 1)$. When the parameters are unknown, we can estimate them by minimizing Q. Bofinger [1] has shown that, in the asymptotic limit, when $N \rightarrow \infty$ while m and the λ's remain fixed, Q has a chi-square distribution. Therefore, parameter estimation by minimizing Q will be called the "minimum chi-square" method.

Unless we want to test goodness of fit between G and the data, the distribution of Q with a finite sample does not matter. We may even take $m = N + 1$ and use spacings of order one, if computational cost is not a concern. Finite-sample properties of the estimator have to be determined by Monte Carlo methods and are beyond the scope of this paper. The important practical point is that Q can be computed even if G is not monotone, and hence no special precautions are needed in the calculations. If G is decreasing in some interval of F in $(0,1)$, that merely worsens fit to the data and increases the value of Q. Experience has shown, however, that minimizing Q may yield small negative slopes at end points. Therefore, it is necessary to determine whether derivatives of the polynomial in equation 1 are nonnegative at $F = 0$ and $F = 1$. If a derivative is negative, it is set equal to zero by changing the coefficients.

Another benefit of using Q is the following. Q is a quadratic function of the coefficients $\underline{a}$. If parameters $\underline{\theta}$ of F are held fixed, minimization over coefficients is achieved by solving simultaneous linear equations. This method is a major simplification of the calculations. In addition, constraining the derivatives at $F = 0$ and $F = 1$ is much easier in linear equations than in nonlinear fitting.

Apart from computational convenience, Q has another advantage over MLE. It is well known that MLE lacks robustness because a single extreme value can dominate the likelihood function and hence the estimates of parameters. In contrast, Q uses not the observed values themselves but their transforms to the probability metric. The transformed value of an observation, no matter how extreme, lies between 0 and 1 and hence cannot dominate the objective function.

ILLUSTRATION

In the Infantry phase of the Marine Corps Job Performance Project, 1,976 Marines were administered a video firing game as a test of eye-hand coordination. The score on this test (rescaled to obtain a mean near 100) was fitted with a Weibull base and a cubic polynomial ($p = 2$) by minimum chi-square. Despite the large sample, the minimized chi-square with 45 degrees of freedom was only 38.7. The parameters of the Weibull were 1.58 and 101.24; the polynomial coefficients were 0.48 and -1.90. Figure 1 shows a histogram and the fitted distribution. Clearly, an excellent fit has been obtained with only four parameters.

APPLICATIONS OF POLYNOMIAL FAMILIES

The primary use of polynomial families, as illustrated above, is to obtain good fit to the sample distribution within the parametric framework. A stepwise fit would be used in most applications. If asymptotic properties of maximum likelihood estimates are to be invoked, the degree of the polynomial needs to be specified in advance.

Polynomial families are also useful for testing goodness of fit. Tests for normality are based on skewness and kurtosis. Corresponding tests can be constructed as follows. If we begin with a normal base and then add only the quadratic term g_1, we obtain a skewed distribution. If the added term is statistically significant, the null hypothesis of normality is rejected in favor of a skewed distribution. Suppose we know or assume that the true distribution is symmetric. Then we can add only the symmetric cubic term $\tilde{g}_1$ (appendix A) and test whether kurtosis is same as that of the normal. We can test skewness and kurtosis simultaneously by adding g_1 and g_2 together.

These tests based on polynomial families have two major advantages over conventional procedures. First, if the null hypothesis is rejected, we can fit an alternative distribution that fits better than the normal one does. Second, the procedure is completely general: it can be used with any base distribution whatsoever (e.g., logistic or Cauchy instead of normal). If the likelihood ratio or chi-square test is used, the asymptotic distribution of the test statistic is the same for all base distributions. (The finite-sample properties of the test statistic will probably depend on the base distribution.)

Thus, polynomial families provide a flexible and hence powerful approach to fitting and testing univariate distributions, discrete as well as continuous.

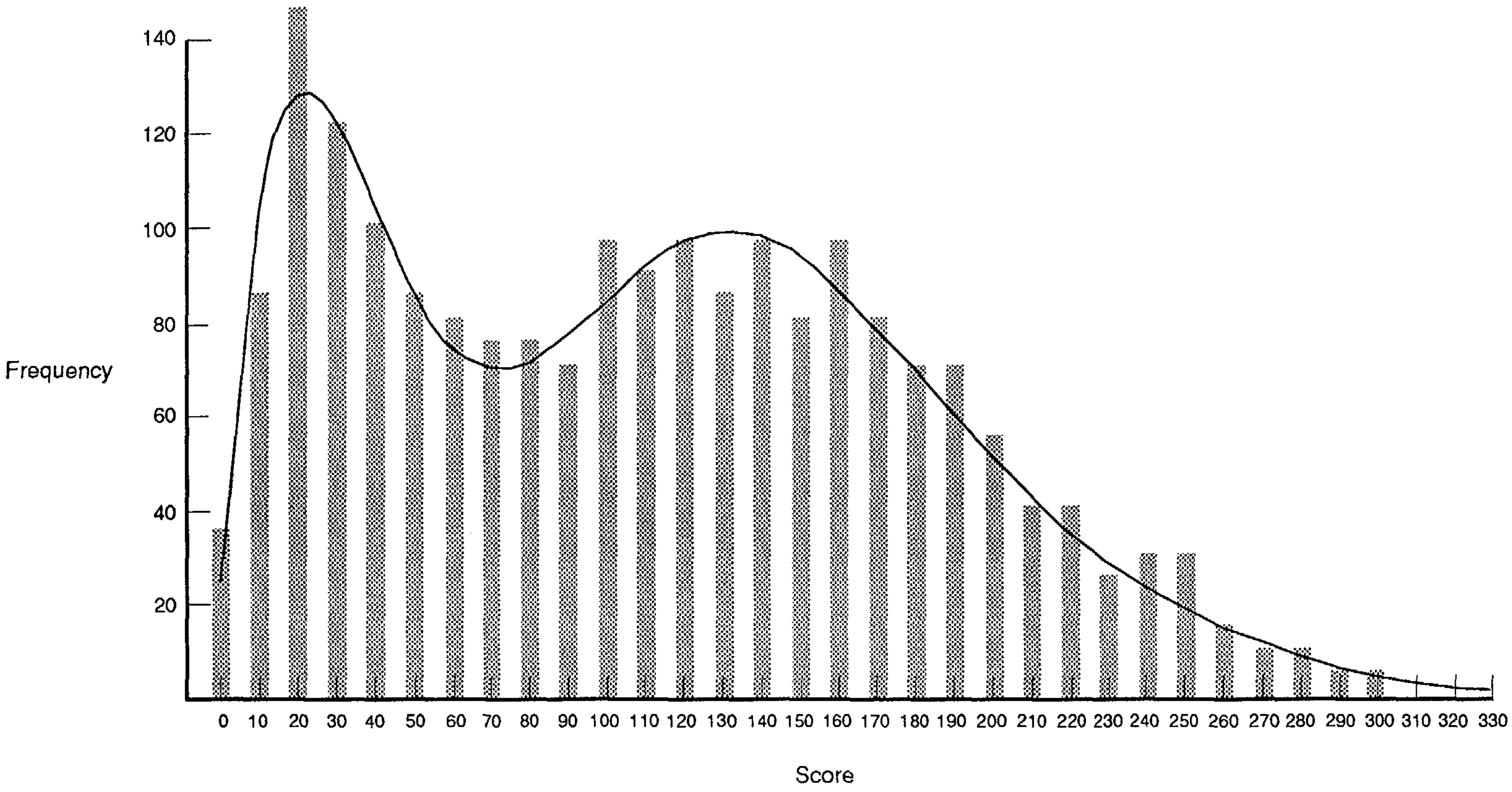

Figure 1. Observed and fitted histograms of video scores

THIS PAGE INTENTIONALLY LEFT BLANK

REFERENCE

[1] Eve Bofinger. "Goodness-of-Fit Test Using Sample Quantiles." *Journal of the Royal Statistical Society,* Series B (1973): 277-284

APPENDIX A

MATHEMATICAL DETAILS FOR CONTINUOUS VARIABLES

g FUNCTIONS AND THEIR DERIVATIVES

The distribution function in the polynomial family is

$$G(x, \underline{\theta}, a) = F(x, \underline{\theta}) + \sum_{k=1}^{p} a_k g_k\{F(x, \underline{\theta})\} \quad ,$$

where F is the base distribution function with parameter vector $\underline{\theta}$.

To simplify computations as well as formulas, we define

$$h = 2F - 1 \ ,$$

$$H = F(1 - F) \ ,$$

$$T = (3 - 16H) \ .$$

Hence,

$$h' = 2 \ ,$$

$$H' = -h \ ,$$

$$T' = 16\ h \ ,$$

where a prime denotes derivative with respect to F.

The functions in the polynomial and their derivatives are

$$g_1 = H \ , \tag{A-1a}$$

$$g_1' = -h \ , \tag{A-1b}$$

$$g_1'' = -2 \ , \tag{A-1c}$$

$$g_1''' = 0 \ ; \tag{A-1d}$$

$$g_2 = F\ H \ , \tag{A-2a}$$

$$g_2' = F(2 - 3F) \ , \tag{A-2b}$$

$$g_2'' = 2 - 6F \ , \tag{A-2c}$$

$$g_2''' = -\ 6 \ ; \tag{A-2d}$$

$$g_3 = H^2 \ , \tag{A-3a}$$

$$g_3' = -2\ h\ H \ , \tag{A-3b}$$

$$g_3'' = 2 - 12H \ , \tag{A-3c}$$

$$g_3''' = 12\ h \ ; \tag{A-3d}$$

$$g_4 = h\ g_3\ , \tag{A-4a}$$

$$g_4' = h\ g_3' + 2\ g_3\ , \tag{A-4b}$$

$$g_4'' = h\ g_3'' + 4\ g_3'\ , \tag{A-4c}$$

$$g_4''' = h\ g_3''' + 6\ g_3''\ ; \tag{A-4d}$$

$$g_5 = T\ g_3\ , \tag{A-5a}$$

$$g_5' = T\ g_3' + 16\ g_4\ , \tag{A-5b}$$

$$g_5'' = T\ g_3'' + 32\ (g_4' - g_3)\ , \tag{A-5c}$$

$$g_5''' = T\ g_3''' + 48\ g_4'' - 96\ g_3'\ ; \tag{A-5d}$$

$$g_6 = h\ g_5\ , \tag{A-6a}$$

$$g_6' = 2\ g_5 + h\ g_5'\ , \tag{A-6b}$$

$$g_6'' = 4\ g_5' + h\ g_5''\ , \tag{A-6c}$$

$$g_6''' = 6\ g_5'' + h\ g_5'''\ ; \tag{A-6d}$$

$$g_7 = T\ g_5\ , \tag{A-7a}$$

$$g_7' = T\, g_5' + 16\, g_6 \;, \qquad \text{(A-7b)}$$

$$g_7'' = T\, g_5'' + 32\, (g_6' - g_5) \;, \qquad \text{(A-7c)}$$

$$g_7''' = T\, g_5''' + 48\, g_6'' - 96\, g_5' \;; \qquad \text{(A-7d)}$$

$$g_8 = h\, g_7 \;, \qquad \text{(A-8a)}$$

$$g_8' = 2\, g_7 + h\, g_7' \;, \qquad \text{(A-8b)}$$

$$g_8'' = 4\, g_7' + h\, g_7'' \;, \qquad \text{(A-8c)}$$

$$g_8''' = 6\, g_7'' + h\, g_7''' \;. \qquad \text{(A-8d)}$$

Apart from g_1, forms of these functions are not unique or even optimal in any sense. They do have a convenient feature: g_1' equals 1 at $F = 0$ and -1 at $F = 1$, and g_2' equals -1 at $F = 1$. All other derivatives vanish at the end points, which makes it easy to impose monotonicity at the end points, where the constraint requires that derivatives not be negative. The corresponding conditions on the coefficients, for $F = 0$ and $F = 1$, are

$$1 + a_1 \geq 0, \text{ i.e., } a_1 \geq -1 \;,$$

and

$$1 - (a_1 + a_2) \geq 0, \text{ i.e., } a_1 + a_2 \leq 1 \;.$$

Let $f(x)$ be the density dF/dx. The fitted density is

$$dG/dx = f\left[1 + \sum_{k=1}^{p} a_k g_k'(F)\right] \quad , \qquad \text{(A-9)}$$

where prime indicates a derivative with respect to F.

FUNCTIONS FOR SYMMETRIC DISTRIBUTIONS

Sometimes it is known (or assumed) that the underlying distribution is symmetric. Let the expression for G be

$$G = F + \sum_{k=1}^{p} \tilde{a}_k \tilde{g}_k(F) \qquad \text{(A-10)}$$

where $\tilde{g}_k$ is of degree $2k + 1$. To ensure that G is a symmetric distribution for all values of the coefficients, the base distribution must be symmetric and each function $\tilde{g}$ in the polynomial must be an odd function of $(F - 1/2)$. Let M be the median of F.

$$\tilde{g}_k\{F(M)\} = \tilde{g}_k(1/2) = 0$$

for all k and hence

$$G(M) = F(M) + 0 = 1/2 \quad .$$

Thus, the density dG/dx is an even function of (x - M), i.e., distribution G is symmetric about its median M.

A convenient choice of polynomials is as follows.

$$\tilde{g}_1 = h\ H \quad , \tag{A-11a}$$

$$\tilde{g}_1' = 6H - 1 \quad , \tag{A-11b}$$

$$\tilde{g}_1'' = -6h \quad , \tag{A-11c}$$

$$\tilde{g}_1''' = -12 \quad ; \tag{A-11d}$$

$$\tilde{g}_2 = H\ \tilde{g}_1 \quad , \tag{A-12a}$$

$$\tilde{g}_2' = H\ (10\ H - 2) \quad , \tag{A-12b}$$

$$\tilde{g}_2'' = 2h - 20\ \tilde{g}_1 \quad , \tag{A-12c}$$

$$\tilde{g}_2''' = 4 - 20\ \tilde{g}_1' \quad ; \tag{A-12d}$$

$$\tilde{g}_3 = T\ \tilde{g}_2 \quad , \tag{A-13a}$$

$$\tilde{g}_3' = T\ \tilde{g}_2' + 16\ h\ \tilde{g}_2 \quad , \tag{A-13b}$$

$$\tilde{g}_3'' = T\ \tilde{g}_2'' + 32(h\ \tilde{g}_2' + \tilde{g}_2) \quad , \tag{A-13c}$$

$$\tilde{g}_3''' = T\, g_2''' + 48\, h\, \tilde{g}_2'' + 96\, \tilde{g}_2' \quad ; \tag{A-13d}$$

$$\tilde{g}_4 = T\, \tilde{g}_3 \quad , \tag{A-14a}$$

$$\tilde{g}_4' = T\, \tilde{g}_3' + 16\, h\, \tilde{g}_3 \tag{A-14b}$$

$$\tilde{g}_4'' = T\, \tilde{g}_3'' + 32(h\, \tilde{g}_3' + \tilde{g}_3) \tag{A-14c}$$

$$\tilde{g}_4''' = T\, \tilde{g}_3''' + 48\, h\, \tilde{g}_3'' + 96\, \tilde{g}_3' \quad . \tag{A-14d}$$

Although written in a different form, functions $\tilde{g}_2$, $\tilde{g}_3$, and $\tilde{g}_4$ are the same as g_4, g_6, and g_8, respectively. The nonnegative slope of the polynomial at $F = 0$ and $F = 1$ requires that

$$\tilde{a}_1 \le 1 \quad .$$

MAXIMUM LIKELIHOOD ESTIMATION

Let $f(x, \underline{\theta})$ be the density function of the base distribution and $\ell(x, \underline{\theta})$ its natural logarithm. The derivative of the fitted distribution function G with respect to F is

$$G'(x, \underline{\theta}, a) = 1 + \sum_k a_k g_k' \{F(x, \underline{\theta})\} \quad . \tag{A-15}$$

Unless otherwise stated, sums over k are from 1 to p and sums over i are from 1 to n. Hence the log likelihood of a sample, containing values x_i, $i = 1, 2, \ldots, n$ is

$$LL = \sum_i \ell(x_i, \underline{\theta}) + \sum_i \log[G'(x_i, \underline{\theta}, \mathbf{a})] \quad . \qquad \text{(A-16)}$$

LL is the function to be maximized. The Newton-Raphson method requires first and second derivatives of LL. Therefore, in estimation by maximum likelihood, we need derivatives of g_k with respect to F up to the third order, and derivatives of F and ℓ with respect to their parameters up to the second order.

Let θ_r denote a parameter of F. For example, if F is a normal distribution, θ_1 is the mean and θ_2 the standard deviation. Subscripts r and s will be used with θ, and subscripts j, k, and l for coefficients and g functions in the polynomial. Subscript i will indicate an observation. Arguments x_i, $\underline{\theta}$, and $\mathbf{a}$ will be suppressed. $\partial F/\partial\theta_r$ and $\partial^2 F/\partial\theta_r\partial\theta_s$ will be abbreviated as $\partial_r F$ and $\partial_r\partial_s F$ and the derivatives of ℓ will be treated similarly. The derivative

$$\partial G'/\partial\theta_r = (\sum_k a_k g_k'') \, \partial F/\partial\theta_r$$

will be denoted by $\partial_r G'$.

The first derivatives of log likelihood are

$$\partial LL/\partial\theta_r = \sum_i [\partial_r \ell + \partial_r G'/G'] \qquad (A\text{-}17)$$

and

$$\partial LL/\partial a_j = \sum_i g_j'/G' \ . \qquad (A\text{-}18)$$

The second derivatives are

$$\partial^2 LL/\partial\theta_r\partial\theta_s = \sum_i [\partial_r\partial_s \ell + (\sum_k a_k g_k'')\ \partial_r\partial_s F/G' - \partial_r G'\ \partial_s G'/G'^2 + \partial_r F\ \partial_s F\ (\sum_k a_k g_k''')\ G'] \ , \qquad (A\text{-}19)$$

$$\partial^2 LL/\partial a_j\ \partial a_l = - \sum_i g_j' g_l'/G'^2 \ , \qquad (A\text{-}20)$$

and

$$\partial LL/\partial\theta_r\ \partial a_j = \sum_i [G' g_j'' - g_j' \sum_k a_k g_k''] \ \partial_r F/G'^2 \ . \qquad (A\text{-}21)$$

Maximizing likelihood using these equations yields simultaneous estimates of $\underline{\theta}$ and the coefficients.

Derivatives for Normal Base

Parameters of the normal distribution are mean θ_1 and standard deviation θ_2. The standard normal variate is

$$z = (x - \theta_1)/\theta_2 \quad .$$

The cdf depends on x only through z:

$$F(x) = \Phi(z)$$

where Φ is the standard normal cdf. Partial derivatives of z with respect to parameters are

$$\partial_1 z = -1/\theta_2 \quad , \tag{A-22a}$$

$$\partial_2 z = -z/\theta_2 \quad , \tag{A-22b}$$

$$\partial_1 \partial_1 z = 0 \quad , \tag{A-22c}$$

$$\partial_2 \partial_2 z = 2z/\theta_2^2 \quad , \tag{A-22d}$$

and

$$\partial_2 \partial_1 z = 1/\theta_2^2 \quad . \tag{A-22e}$$

Derivatives of Φ with respect to z are

$$\Phi' = \phi = \exp(-z^2/2)/\sqrt{2\pi} ,$$

which is the standard normal density, and

$$\Phi'' = \phi' = -z\phi .$$

Using these derivatives of z and of Φ, those of $F(x)$ can be computed as follows:

$$\partial_1 F = \phi \partial_1 z = - \phi/\theta_2 , \tag{A-23a}$$

$$\partial_2 F = \phi \partial_2 z = -z\phi/\theta_2 , \tag{A-23b}$$

$$\partial_1 \partial_1 F = -z\phi/\theta_2^2 , \tag{A-23c}$$

$$\partial_2 \partial_2 F = \phi' z^2/\theta_2^2 + 2z\phi/\theta_2^2 ,$$

$$= z\phi(2 - z^2) , \tag{A-23d}$$

and

$$\partial_2 \partial_1 F = \phi(1 - z^2)/\theta_2^2 . . \tag{A-23e}$$

The density of x is

$$f(x, \underline{\theta}) = \phi(z)/\theta_2$$

and hence the log density is

$$\ell(x, \underline{\theta}) = - z^2/2 - \log(\theta_2) - [\log(2\pi)]/2 \quad .$$

Therefore, its partial derivatives are

$$\partial_1\ell = -z\ \partial_1 z = z/\theta_2 \qquad \text{(A-24a)}$$

$$\partial_2\ell = -z\ \partial_2 z - 1/\theta_2$$

$$= z^2/\theta_2 - 1/\theta_2 \quad , \qquad \text{(A-24b)}$$

$$\partial_1\partial_1\ell = (\partial_1 z)/\theta_2 = -1/\theta_2^2 \quad , \qquad \text{(A-24c)}$$

$$\partial_2\partial_2\ell = 2z\ \partial_2 z/\theta_2 - z^2/\theta_2^2 + 1/\theta_2^2$$
$$= -3z^2/\theta_2^2 + 1/\theta_2^2 \quad , \qquad \text{(A-24d)}$$

$$\partial_2\partial_1\ell = \partial_2 z/\theta_2 - z/\theta_2^2$$
$$= -2z/\theta_2^2 \quad . \qquad \text{(A-24e)}$$

Derivatives for Weibull Base

The Weibull cdf is

$$F\ (x,\ \underline{\theta}) = 1 - \exp(-z) \qquad \text{(A-25a)}$$

where

$$z = (x/\theta_2)^{\theta_1}\ . \qquad \text{(A-25b)}$$

Thus, θ_1 and θ_2 are shape and scale parameters. Derivatives of z are

$$\partial_1 z = z \log(x/\theta_2)\ , \qquad \text{(A-26a)}$$

$$\partial_2 z = -\theta_1 z/\theta_2\ , \qquad \text{(A-26b)}$$

$$\partial_1\partial_1 z = \log(x/\theta_2)\ \partial_1 z\ , \qquad \text{(A-26c)}$$

$$\begin{aligned}\partial_2\partial_2 z &= -\theta_1\ \partial_2 z/\theta_2 + \theta_1\ z/\theta_2^2 \\ &= -(\theta_1+1)\ \partial_2 z/\theta_2\ , \end{aligned} \qquad \text{(A-26d)}$$

$$\begin{aligned}\partial_2\partial_1 z &= \partial_2 z\ \log(x/\theta_2)\ -z/\theta_2 \\ &= \{\log(x/\theta_2) + 1/\theta_1\}\partial_2 z\ .\end{aligned} \qquad \text{(A-26e)}$$

Derivatives of F with respect to parameters are

$$\partial_1 F = (1 - F)\ \partial_1 z \quad , \tag{A-27a}$$

$$\partial_2 F = (1 - F)\ \partial_2 z \quad , \tag{A-27b}$$

$$\partial_1 \partial_1 F = (1 - F)\ \partial_1 z\ [\log(x/\theta_2) - \partial_1 z] \quad , \tag{A-27c}$$

$$\partial_2 \partial_2 F = -(1 - F)\ [(\theta_1 + 1)/\theta_2 + \partial_2 z]\ \partial_2 z \quad , \tag{A-27d}$$

and

$$\partial_2 \partial_1 F = (1 - F)[1/\theta_1 - \partial_1 z + \log(x/\theta_2)]\partial_2 z \quad . \tag{A-27e}$$

The density function is

$$f(x, \underline{\theta}) = \theta_1\ (x/\theta_2)^{\theta_1 - 1} \exp(-z)/\theta_2 \quad ,$$

and its log is

$$\ell(x, \underline{\theta}) = \log(\theta_1) + (\theta_1 - 1) \log(x) - \theta_1 \log(\theta_2) - z \quad ,$$

which has derivatives

$$\partial_1 \ell = 1/\theta_1 + \log(x/\theta_2) - \partial_1 z \quad , \qquad \text{(A-28a)}$$

$$\partial_2 \ell = -\theta_1/\theta_2 - \partial_2 z = \theta_1 (z - 1)/\theta_2 \quad , \qquad \text{(A-28b)}$$

$$\partial_1 \partial_1 \ell = -1/\theta_1^2 - \partial_1 \partial_1 z \quad , \qquad \text{(A-28c)}$$

$$\partial_2 \partial_2 \ell = \theta_1/\theta_2^2 - \partial_2 \partial_2 z \quad , \qquad \text{(A-28d)}$$

and

$$\partial_2 \partial_1 \ell = -1/\theta_2 - \partial_2 \partial_1 z \quad . \qquad \text{(A-28e)}$$

Estimation by Minimum Chi-Square

The quantity Q to be minimized, defined in equation 4 in the main text, is

$$Q = \sum_h [(N + 1)\{G(x_{(N_h)} , \underline{\theta}, a) - G(x_{(N_{h-1})}, \underline{\theta}, a)\} - n_h]^2/n_h \quad ,$$

which simplifies to

$$Q = (N + 1)^2 \sum_h \{G(x_{(N_h)}, \underline{\theta}, a) - G(x_{(N_{h-1})}, \underline{\theta}, a)]^2/n_h - (N + 1) \quad .$$

Sums over h are from 1 to m.

Denote $F(x_{(N_h)}, \underline{\theta})$ by F_h and $G(x_{(N_h)}, \underline{\theta}, a)$ by G_h. Then

$$Q = (N + 1)^2 \sum_h [(F_h - F_{h-1}) + \sum_k a_k\{g_k(F_h) - g_k(F_{h-1})\}]^2/n_h - (N + 1) \quad . \text{ (A-29)}$$

First, let us minimize Q over the coefficients **a** while the parameters $\underline{\theta}$ of F are fixed:

$$\partial Q/\partial a_j = 2\ (N + 1)^2 \sum_h [(F_h - F_{h-1}) + \sum_k a_k\ \{g_k(F_h) - g_k(F_{h-1})\}]$$
$$\{g_j(F_h) - g_j(F_{h-1})\}/n_h \quad .$$

Setting these derivatives equal to zero yields simultaneous linear equations of the form

$$\sum_k C_{jk}\ a_k = c_j \quad , \qquad \text{(A-30a)}$$

where

$$c_j = -\sum_h (F_h - F_{h-1})\ \{g_j(F_h) - g_j(F_{h-1})\}/n_h \qquad \text{(A-30b)}$$

and

$$C_{jk} = C_{kj} = \sum_h \{g_j(F_h) - g_j(F_{h-1})\}\ \{g_k(F_h) - g_k(F_{h-1})\}/n_h \quad . \qquad \text{(A-30c)}$$

To obtain derivatives of the coefficients with respect to $\underline{\theta}$, we differentiate equation A-30b and rearrange terms to obtain

$$\sum_k C_{jk}\,\partial_r a_k = \partial_r c_j - \sum_k \partial_r C_{jk}\,a_k \tag{A-31a}$$

and

$$\begin{aligned}\sum_k C_{jk}\,\partial_r\partial_s a_k = \partial_r\partial_s c_j &- \sum_k \partial_s C_{jk}\,\partial_r a_k \\ &- \sum_k \partial_r C_{jk}\,\partial_s a_k - \sum_k \partial_r\partial_s C_{jk}\,a_k \quad .\end{aligned} \tag{A-31b}$$

Derivatives needed in these equations are

$$\begin{aligned}\partial_r c_j = -\sum_h \big[&(\partial_r F_h - \partial_r F_{h-1})\ \{g_j(F_h) - g_j(F_{h-1})\} \\ &+ (F_h - F_{h-1})\ \{\partial_r F_h\ g_j'(F_h) - \partial_r F_{h-1}\ g_j'(F_{h-1})\}\big]/n_h \quad ,\end{aligned} \tag{A-32a}$$

$$\begin{aligned}\partial_r\partial_s c_j = -\sum_h \big[&(\partial_r\partial_s F_h - \partial_r\partial_s F_{h-1})\ \{(g_j(F_h) - g_j(F_{h-1})\} \\ &+ (\partial_r F_h - \partial_r F_{h-1})\ \{\partial_s F_h\ \ g_j'(F_h) - \partial_s F_{h-1}\ g_j'(F_{h-1})\} \\ &+ (\partial_s F_h - \partial_s F_{h-1})\ \ \{\partial_r F_h\ g_j'(F_h) - \partial_r F_{h-1}\ g_j'(F_{h-1})\} \\ &+ (F_h - F_{h-1})\ \{\partial_r\partial_s F_h\ g_j'(F_h) + \partial_r F_h\ \partial_s F_h\ g_j''\ (F_h) \\ &-\partial_r\partial_s F_{h-1}\ g_j'(F_{h-1}) - \partial_r F_{h-1}\ \ \partial_s F_{h-1}\ g_j''\ (F_{h-1})\}\big]/n_h \ ,\end{aligned} \tag{A-32b}$$

$$\begin{aligned}\partial_r C_{jk} = \sum_h \big[&\{\partial_r F_h\ g_j'(F_h) - \partial_r F_{h-1}\ g_j'(F_{h-1})\}\{g_k(F_h) - g_k(F_{h-1})\} \\ &+ \{g_j(F_h) - g_j(F_{h-1})\}\{\partial_r F_h\ g_k'(F_h) - \partial_r F_{h-1}\ g_k'(F_{h-1})\}\big]/n_h \quad ,\end{aligned} \tag{A-33a}$$

$$\begin{aligned}\partial_r\partial_s C_{jk} = \sum_h \big[&\{\partial_r\partial_s F_h\ \ g_j'(F_h) + \partial_r F_h\ \partial_s F_h\ g_j''(F_h) \\ &- \partial_r\partial_s F_{h-1}\ g_j'(F_{h-1}) - \partial_r F_{h-1}\ \partial_s F_{h-1}\ g_j''(F_{h-1})\}\{g_k(F_h) - g_k\ (F_{h-1})\}\end{aligned}$$

$$\begin{aligned}
&+ \{\partial_r F_h\, g_j'(F_h) - \partial_r F_{h-1}\, g_j'(F_{h-1})\}\ \{\partial_s F_h\, g_k'(F_h) - \partial_s F_{h-1}\, g_k'(F_{h-1})\} \\
&+ \{\partial_s F_h\, g_j'(F_h) - \partial_s F_{h-1}\, g_j'(F_{h-1})\}\ \{\partial_r F_h\, g_k'(F_h) - \partial_r F_{h-1}\, g_k'(F_{h-1})\} \\
&+ \{g_j(F_h) - g_j(F_{h-1})\}\ \{\partial_r \partial_s F_h\, g_k'(F_h) + \partial_r F_h\, \partial_s F_h\, g_k''(F_{h-1}) \\
&- \partial_r \partial_s F_{h-1}\, g_k'(F_{h-1}) - \partial_r F_{h-1}\, \partial_s F_{h-1}\, g_k''(F_{h-1})\}]/n_h \quad . \qquad \text{(A-33b)}
\end{aligned}$$

At any given $\underline{\theta}$, one computes the coefficients in the polynomial and then the sum of squares Q. Thus, the coefficients and Q are functions of $\underline{\theta}$. The derivatives of Q are given by

$$\begin{aligned}
\{1/2(N+1)^2\}\partial_r Q = {} & \sum_h (G_h - G_{h-1})[(\partial_r F_h - \partial_r F_{h-1}) + \sum_k a_k\{\partial_r F_h\, g_k'(F_h) \\
& -\partial_r F_{h-1}\, g_k'(F_{h-1})\} + \sum_k \partial_r a_k\ \{g_k(F_h) - g_k(F_{h-1})\}]/n_h \ , \qquad \text{(A-34a)}
\end{aligned}$$

$$\begin{aligned}
[1/2(N + 1)^2]\ \partial_r \partial_s Q/ = {} & \sum_h [(\partial_r G_h - \partial_r G_{h-1})\ (\partial_s G_h - \partial_s G_{h-1}) \\
& + (G_h - G_{h-1})\ (\partial_r \partial_s F_h - \partial_r \partial_s F_{h-1})]/n_h \\
& + \sum_h (G_h - G_{h-1})[\ \sum_k a_k\{\partial_r \partial_s F_h\, g_k'(F_h) \\
& + \partial_r F_h\, \partial_s F_h\, g_k''(F_h) - \partial_r \partial_s F_{h-1}\, g_k'(F_{h-1}) \\
& - \partial_r F_{h-1}\, \partial_s F_{h-1}\, g_k''(F_{h-1})\} + \sum_k \partial_r a_k\ \{\partial_s F_h\, g_k'(F_h) \\
& - \partial_s F_{h-1}\, g_k'(F_{h-1})\} + \sum_k \partial_r \partial_s a_k\ \{g_k(F_h) - g_k(F_{h-1})\} \\
& + \sum_k \partial_s a_k\{\partial_r F_h\, g_k'(F_h) - \partial_r F_{h-1} g_k'(F_{h-1})\}]/n_h \quad .
\end{aligned}$$

(A-34b)

where $(\partial_r G_h - \partial_r G_{h-1})$ is the quantity in square brackets in equation A-34a, and $(\partial_s G_h - \partial_s G_{h-1})$ is defined similarly. These derivatives are used to estimate $\underline{\theta}$ by minimizing Q with the Newton-Raphson procedure.

Constraint at F = 0

The derivative of G with respect to F must be nonnegative at F = 0. This requires that $a_1 \geq -1$. If we obtain $a_1 < -1$ after solving the linear equations, the value obtained is replaced by -1. If p = 1, (i.e., if no terms higher than the quadratic are present), we set $\partial_r a_1 = \partial_r \partial_s a_1 = 0$ for all r and s and then proceed to calculate Q and its derivatives.

If p > 1, the polynomial is reexpressed in the form

$$G = F - g_1(F) + \sum_{k=1}^{p-1} b_k e_k(F) \quad , \tag{A-35a}$$

where

$$e_k(F) = g_{k+1}(F) \quad . \tag{A-35b}$$

Equations for the coefficients **b** are of the form

$$\sum_{k=1}^{p-1} D_{jk} b_k = d_j \quad , \tag{A-36a}$$

where

$$D_{jk} = C_{j+1,\ k+1} \qquad \text{(A-36b)}$$

and

$$d_j = c_{j+1} + C_{1,j+1} \quad . \qquad \text{(A-36c)}$$

The derivatives of **D** and **d**, and hence those of **b**, can be computed from those of **C** and **c**. Then we have $a_1 = -1$, $\partial_r a_1 = \partial_r \partial_s a_1 = 0$ and, for $j > 1$ $a_j = b_{j-1}$, $\partial_r a_j = \partial_r b_{j-1}$, and $\partial_r \partial_s a_j = \partial_r \partial_s b_{j-1}$.

For symmetric distributions, the constraint is $\tilde{a}_1 < 1$. Equations for imposing this constraint are similar to those above for the general case, with $\tilde{a}_1 = +1$ and $-g_1$ replaced by $+\tilde{g}_1$ in equation A-35a, and a negative sign for $C_{1,j+1}$ in equation A-36c. Because of symmetry, derivatives at $F = 0$ and $F = 1$ are equal and hence the derivative at $F = 1$ need not be checked separately.

Constraint at F = 1

The derivative of G with respect to F at $F = 1$ must be nonnegative, which requires $a_1 + a_2 \geq 1$. If this condition is violated, we write

$$G = F + g_1 + a_2(g_2 - g_1) + \sum_{k=3}^{p} a_k g_k$$

$$= F + g_1 + \sum_{k=1}^{p-1} b_k e_k \quad , \qquad \text{(A-37a)}$$

where

$$e_1 = g_2 - g_1 \quad , \qquad \text{(A-37b)}$$

$$e_k = g_{k+1} \text{ if } k > 1 \quad , \qquad \text{(A-37c)}$$

and

$$b_k = a_{k+1} \quad . \qquad \text{(A-37d)}$$

Equations for b_k have the form in equation A-36a with

$$D_{jk} = C_{j+1,k+1} \text{ if } j, k > 1 \quad , \qquad \text{(A-38a)}$$

$$D_{1k} = C_{2,k+1} - C_{1,k+1} \text{ if } k > 1 \quad , \qquad \text{(A-38b)}$$

$$D_{11} = C_{22} + C_{11} - 2C_{12} \quad , \qquad \text{(A-38c)}$$

$$d_k = c_{k+1} - C_{1,k+1} \text{ if } k > 1 \quad , \qquad \text{(A-38d)}$$

and

$$d_1 = c_2 - c_1 - (C_{12} - C_{11}) \quad . \tag{A-38e}$$

After coefficients **b** and their derivatives have been computed, the original coefficients are obtained as

$$a_1 = 1 - b_1 \ , \tag{A-39a}$$

$$a_k = b_{k-1} \text{ if } k > 1 \quad . \tag{A-39b}$$

and similarly for their derivatives.

Constraints at F = 0 and F = 1

If a zero slope has to be imposed at both end points, we write

$$G = F - g_1 + 2g_2 + \sum_{k=1}^{p-2} b_k e_k \quad , \tag{A-40a}$$

where

$$e_k = g_{k+2} \quad . \tag{A-40b}$$

The coefficients are obtained by solving

$$\sum_{k=1}^{p-2} D_{jk} b_k = d_j \tag{A-41a}$$

APPENDIX B

MATHEMATICAL DETAILS FOR DISTRIBUTION OF TEST SCORES

INTRODUCTION

Consider a test containing n items. The test score x is the number of items answered correctly, so that $0 \le x \le n$. A convenient base distribution for test scores is the beta binomial distribution generated as follows. Let T have a beta distribution with parameters a and b'. Conditional on $T = t$, let the distribution of x be binomial with mean nt. Then, integration over t shows that the marginal probability of score x is

$$f(x) = [\Gamma(n + 1)/\Gamma(x + 1)\Gamma(n - x + 1)]\ [\Gamma(x + a)\Gamma(n - x + b')/\Gamma(n + a + b')]$$
$$[\Gamma(a + b')/\Gamma(a)\Gamma(b')] \tag{B-1}$$

This is a special case of the hypergeometric distribution, called the negative hypergeometric.

Following Lord and Novick [B-1, section 23.6], it is convenient to replace parameter b' with

$$b = b' + n - 1 \quad .$$

The parameters of the distribution can be calculated from the mean μ and standard deviation σ as follows:

$$a = \mu(-1 + 1/\alpha) \qquad \text{(B-2a)}$$

and

$$b = -a - 1 + n/\alpha \quad , \qquad \text{(B-2b)}$$

where

$$\alpha = [n/(n - 1)] \; [1 - \mu(n - \mu)/n\sigma^2] \quad . \qquad \text{(B-2c)}$$

Estimates of a and b can be obtained by replacing μ and σ with the sample mean and standard deviation.

CALCULATION OF PROBABILITIES

The score probabilities can be calculated by recursion. On rewriting equation 23.6.4 in [B-1], the ratio of the probabilities of successive scores is given by

$$f(x + 1)/f(x) = (n - x) \; (a + x)/(x + 1) \; (b - x), \; x \le n - 1 \quad . \qquad \text{(B-3)}$$

Denote this ratio by $w(x)$. Let u be some score in the middle of the distribution, say the largest integer smaller than the mean. Let $v(u) = 1$.

$$v(x + 1) = w(x)v(x) \ , \ u \leq x \leq n - 1 \qquad \text{(B-4a)}$$

and

$$v(x) = v(x + 1)/w(x) \ , \ 0 \leq x < u \ . \qquad \text{(B-4b)}$$

Then the score probabilities are given by

$$f(x) = v(x)/ \sum_{y=0}^{n} v(y) \qquad \text{(B-5a)}$$

and then the cumulative probabilities by

$$F(x) = \sum_{y=0}^{x} f(y) \ . \qquad \text{(B-5b)}$$

DERIVATIVES OF PROBABILITIES

The parameters of the distribution are $\theta_1 = a$ and $\theta_2 = b$. Therefore, using the same abbreviations for derivatives as in appendix A,

$$\partial_1 w(x) = w(x)/(a + x) \ , \qquad \text{(B-6a)}$$

$$\partial_2 w(x) = -w(x)/(b - x) \quad , \tag{B-6b}$$

$$\partial_1 \partial_1 w(x) = 0 \quad , \tag{B-6c}$$

$$\partial_2 \partial_2 w(x) = 2\ w(x)/(b - x)^2 \quad , \tag{B-6d}$$

and

$$\partial_1 \partial_2 w(x) = -w(x)/(a + x)\ (b - x) \quad . \tag{B-6e}$$

All derivatives of $v(u)$ vanish. For $u \le x < n$, we have

$$\partial_r\ v(x + 1) = w(x)\ \partial_r v(x) + v(x)\ \partial_r w(x) \tag{B-7a}$$

and

$$\begin{aligned} \partial_r \partial_s v(x + 1) = {} & w(x)\ \partial_r \partial_s v(x) + \partial_r w(x) \partial_s v(x) + \partial_s w(x) \partial_r v(x) \\ & + v(x)\ \partial_r \partial_s w(x) \quad , \end{aligned} \tag{B-7b}$$

where each subscript can be 1 or 2. By rearranging equations B-7a and B-7b, corresponding equations useful at $x < u$ are found to be

$$\partial_r v(x) = [\partial_r v(x + 1) - v(x) \partial_r w(x)]/w(x) \quad , \tag{B-8a}$$

and

$$\partial_r\partial_s v(x) = [\partial_r\partial_s v(x+1) - \partial_r w(x)\partial_s v(x) - \partial_s w(x)\partial_r v(x) - v(x)\partial_r\partial_s w(x)]/w(x) \quad . \tag{B-8b}$$

where derivatives of $w(x)$ are obtained from equations B-6a to B-6e.

To obtain derivatives of score probabilities, define the sum

$$S = \sum_{x=0}^{n} v(x) \quad ,$$

and use equation B-5a to obtain

$$\partial_r f(x) = [\partial_r v(x) - f(x)\partial_r S]/S \tag{B-9a}$$

and

$$\partial_r\partial_s f(x) = [\partial_r\partial_s v(x) - f(x)(\partial_r\partial_s S) - \partial_s f(x)(\partial_r S) - \partial_r f(x)(\partial_s S)]/S \tag{B-9b}$$

ESTIMATION

The fitted cdf is

$$G(x, \underline{\theta}, \mathbf{a}) = F(x) + \sum_k a_k g_k\, [F(x, \underline{\theta})] \quad . \tag{B-10}$$

Hence the fitted score probabilities are given by

$$\mathrm{pr}(0, \underline{\theta}, a) = F(0, \underline{\theta}) + \sum_k a_k g_k[F(0, \underline{\theta})] \qquad \text{(B-11a)}$$

and, for $x > 0$,

$$\mathrm{pr}(x, \underline{\theta}, a) = f(x, \underline{\theta}) + \sum_k a_k \; [g_k\{F(x, \underline{\theta})\}$$

$$- g_k\{F(x - 1, \underline{\theta})\}] \qquad \text{(B-11b)}$$

These probabilities can be used for maximum likelihood estimation, but minimum chi-square is more convenient while being asymptotically equivalent to maximum likelihood.

For minimum chi-square estimation, create $m \leq n + 1$ cells or score groups by choosing scores $0 \leq y_1 < y_2 \ldots < y_m = n$ so that the observed frequency in each score group exceeds some value (say 10). Let $F_h = F(y_h, \underline{\theta})$ and $G_h = G(y_h, \underline{\theta}, a)$, with $y_0 = -1$ and $F_0 = G_0 = 0$ by definition, and $F_m = G_m = 1$. Let n_h be the observed frequency in cell h which contains scores x given by $y_{h-1} + 1 \leq x \leq y_h$, so that $\sum_h n_h = N$, the sample size. The quantity to be minimized is

$$Q = \sum_h [N\,(G_h - G_{h-1}) - n_h]^2/n_h \quad . \qquad \text{(B-12)}$$

Expressions for derivatives of Q are the same as in appendix A, except that $(N + 1)$ (which in appendix A is the total number of gaps) is replaced by the sample size N.

REFERENCE

[1] Frederic M. Lord and Melvin R. Novick. *Statistical Theories of Mental Test Scores*. Reading, Mass: Addison Wesley, 1968